Department of Defense
Sexually Transmitted Infections:
Estimation of Burden among Active Duty Service Members using Clinical Diagnoses, Laboratory Results, and Medical Event Reports

Global Emerging Infections Surveillance and Response System (GEIS) Project, Fiscal Year 2008

Ashleigh K. McCabe, MPH
EpiData Center Department
Communicable Disease Division

Updated: March 2016
NMCPHC-TR-EDC-191-2016

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
	Technical Report	October 2007 - September 2008

4. TITLE AND SUBTITLE
Department of Defense Sexually Transmitted Infections:
Estimation of Burden among Active Duty Service Members using Clinical Diagnoses, Laboratory Results, and Medical Event Reports

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ashleigh McCabe

5d. PROJECT NUMBER
GP0033_08_NE

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy and Marine Corps Public Health Center
EpiData Center Department
620 John Paul Jones Circle, Suite 1100
Portsmouth, VA 23708

8. PERFORMING ORGANIZATION REPORT NUMBER
NMCPHC-TR-EDC-191-2016

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DOD Global Emerging Infections Surveillance and Response System
Armed Forces Health Surveillance Branch
503 Robert Grant Avenue
Silver Spring, MD 20910

10. SPONSOR/MONITOR'S ACRONYM(S)
DOD GEIS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release. Distribution is not limited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project established the foundation for the use of laboratory and encounter data in sexually transmitted infection (STI) surveillance and applied the capture-recapture statistical methodology to estimate cases not captured in laboratory results, clinical diagnoses, or medical event reports (MERs). Methods used to extract reportable STIs (chlamydia, gonorrhea, syphilis) in CY 2006 across the Department of Defense are detailed in this report. Cases from the three sources were then stacked and labeled with which database(s) it was identified using a 30-day gap-in-care rule. When data were matched, the majority of cases for each disease were from only one database. There were few cases identified in all three sources: 1.0% of chlamydia, 5.0% of gonorrhea, and no syphilis cases. The capture-recapture statistical method estimated that approximately 154 syphilis cases (95% CI: 147.6, 160.4) may be missing when using the three data sources, for a total estimated case burden of 356 cases. Though using all three sources for surveillance may require more resources than single-source surveillance methods, using only one source does not provide a complete picture of true STI burden. Using diagnoses or MERs alone would have identified less than a tenth of chlamydia cases and a fourth of gonorrhea cases.

15. SUBJECT TERMS
Sexually Transmitted Infection (STI), Chlamydia, Gonorrhea, Syphilis, Active Duty, Case Definition, Burden, Capture-Recapture

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ashleigh McCabe
U	U	U	UU	21	19b. TELEPHONE NUMBER *(Include area code)* 757.953.0692

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Abstract

This project established the foundation for the use of laboratory and encounter data in sexually transmitted infection (STI) surveillance and applied the capture-recapture statistical methodology to estimate cases not captured in laboratory results, clinical diagnoses, or medical event reports (MERs). Methods used to extract reportable STIs (chlamydia, gonorrhea, syphilis) in CY 2006 across the Department of Defense are detailed in this report. Cases from the three sources were then stacked and labeled with which database(s) it was identified using a 30-day gap-in-care rule. When data were matched, the majority of cases for each disease were from only one database. There were few cases identified in all three sources: 1.0% of chlamydia, 5.0% of gonorrhea, and no syphilis cases. The capture-recapture statistical method estimated that approximately 154 syphilis cases (95% CI: 147.6, 160.4) may be missing when using the three data sources, for a total estimated case burden of 356 cases. Though using all three sources for surveillance may require more resources than single-source surveillance methods, using only one source does not provide a complete picture of true STI burden. Using diagnoses or MERs alone would have identified less than a tenth of chlamydia cases and a fourth of gonorrhea cases.

Table of Contents

Abstract ii
Executive Summary 1
Introduction 1
Methods 1
Results 1
Discussion 1
Introduction 2
Methods 3
Basic Methodology 3
Case Identification 3
Case Definition 3
Data Processing 6
Removing Duplicate Records 6
DMSS Records 6
Encounter Records 6
Laboratory Data 6
Matching 8
Results 8
Cases 8
Laboratory Data 8
Encounter Records 9
DMSS Records 9
Demographic Comparisons 9
Chlamydia 9
Gonorrhea 10
Syphilis 11
Matching 11
Capture-Recapture 12
Discussion 13
Acronym List 16

List of Figures and Tables

Table 1. Capture-Recapture Methodology Matrix....3
Table 2. STI Case Definitions....3
Table 3. ICD-9-CM Codes for Identification of Cases....4
Table 4. Dataset Search Terms....5
Table 5. Number of Unique Cases Identified in each Database by Infection for DOD Active Duty Personnel in CY2006....8
Table 6. Number of DOD Active Duty Service Members by Infection in CY2006....8
Table 7. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Chlamydia Infection in CY2006....10
Table 8. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Gonorrhea Infection in CY2006....11
Table 9. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Syphilis Infection in CY2006....11
Table 10. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Chlamydia....12
Table 11. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Gonorrhea....12
Table 12. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Syphilis....12

Executive Summary

Introduction

To estimate the Department of Defense (DOD) Sexually Transmitted Infection (STI) burden, the EpiData Center (EDC) applied the capture-recapture statistical methodology to evaluate the relationship between medical event reports, laboratory results, and administrative/clinical encounter databases. Efforts in fiscal year 2008 (FY2008) established the foundation for the use of the laboratory and encounter databases to aid in the estimation of reportable STI burden on the military and enhance infection surveillance activities.

Methods

This project focused on reportable STIs (chlamydia, gonorrhea, and syphilis) in calendar year 2006 (CY2006) from all DOD military treatment facilities (MTFs). The Defense Medical Surveillance System (DMSS) was used to capture medical event reports, the Military Health System (MHS) Management Analysis and Reporting Tool (M2) was queried for inpatient and outpatient clinical encounter records. And laboratory records were extracted from Health Level 7 (HL7) formatted chemistry and microbiology datasets. Methods used to extract cases from each data source are detailed in this report. Cases from the three sources were then stacked labeled with which database(s) it was identified using a 30-day gap-in-care rule. The capture-recapture methodology was applied to syphilis results to assess the number of cases not captured in any electronic data source used.

Results

The majority (69.9%) of cases were identified in the HL7 formatted laboratory records. There were fewer cases identified in the encounter and DMSS databases with 7.8% and 22.3%, respectively. When data were matched, similar characteristics were observed among all infections, with most records present in only one database. There were few cases identified in all three sources: 1.0% of chlamydia cases, 5.0% of gonorrhea cases, and no syphilis cases. The capture-recapture statistical method was applied to the case distribution for syphilis. The model of best fit estimated that approximately 154 cases (95% CI: 147.6, 160.4) may be missing when using the three data sources, for a total estimated case burden of 356 cases.

Discussion

Though using all three sources for surveillance may require more resources than single-source surveillance methods, using only one source does not provide a complete picture of the true STI burden. This project found that a vast majority (81.8%) of chlamydia and gonorrhea cases were identified in the laboratory data, the remainder being found in encounter and/or reporting records. Using encounter or reporting databases alone would have only identified less than a tenth of chlamydia cases and a fourth of gonorrhea cases. The most significant advantage of using laboratory records for case identification is that laboratory results are often definitive. Positive laboratory testing continues to be the gold standard for surveillance systems and case finding projects.

Introduction

DOD medical surveillance tracks specific medical infections and environmental exposures of military interest. An estimate of the STI burden in the DOD is necessary to target intervention and preventive medicine programs to the populations with the highest need. An estimate allows for comparison to the general civilian population to determine the risk and impact that STIs have on the DOD military population, troop strength, and readiness. Efforts have been made to estimate STI burden in DOD populations, however, early attempts focused on specific subpopulations and because they were performed strictly through the use of a single data source were likely underestimates of the true burden. Furthermore, previous burden estimates were primarily based on symptomatic cases of STIs, and did not account for asymptomatic cases that contribute to the infection burden. Estimation of STIs in the Navy and Marine Corps has the added complexity of non-shore based facilities whose records are not included in common clinical data sources.

The EDC Department of the Navy and Marine Corps Public Health Center (NMCPHC) conducted a pilot study to evaluate the existing relationship between medical event reports in the Naval Disease Reporting System-internet version (NDRSi), HL7 formatted laboratory results extracted from the Composite Health Care System (CHCS), and inpatient and outpatient clinical encounter data [Standard Ambulatory Data Record (SADR), Standard Inpatient Data Record (SIDR)] using the capture-recapture statistical method. The pilot study was performed on data from a single shore-based command with some impact from operational forces and a transient population. The evaluation extracted chlamydia cases from each of the databases based on the Tri-Service Reportable Events Guidelines and Case Definitions (May 2004). The databases were aligned to assess the number of cases identified in all or any combination of sources. Based on the initial methodology developed for the pilot study, the EDC is confident that the methodology can be utilized to provide a more precise estimate of infection burden in the DOD.

To estimate the STI burden in the DOD, in FY2008 the EDC applied the lessons learned from the pilot study and the capture-recapture statistical methodology to evaluate the relationship between medical event reports, laboratory results, and administrative/clinical encounter databases. Using the case definitions in the Tri-Service Reportable Events Guidelines and Case Definitions (May 2004), the methodology estimates the cases not captured by any of the databases. For the purposes of this project, only reportable STIs (chlamydia, gonorrhea, and syphilis) in CY2006 were considered. Efforts in FY2008 establish the foundation for the use of the laboratory and diagnostic databases to aid in the estimation of reportable STI burden on the military and infection surveillance activities. While the current directives require reporting all cases, compliance is inconsistent across the various MTFs and/or commands.

The DOD Global Emerging Infections Surveillance and Response System (GEIS) funded this project during FY2008.

Methods

For the case identification portion of this project, each database was considered separately due to the differences in data structure. The databases used differed from those of the pilot study as not all original datasets contained records for the entire DOD. DMSS was used instead of DRSi to capture medical event reports from the entire DOD. M2 was queried for inpatient and outpatient clinical encounter records. As in the pilot study, laboratory records were extracted from HL7 formatted chemistry and microbiology datasets. The study included all DOD MTFs and assessed chlamydia, gonorrhea, and syphilis.

Basic Methodology

1. Each database was analyzed and divided into three subgroups, one for each STI type: chlamydia, gonorrhea, and syphilis.
2. By infection, cases found in each database were matched to each other. Database alignment was assessed.
3. After matching was completed, the capture-recapture methodology was applied to estimate the number of cases not captured in any of the datasets. The overall goal of the study was to complete the following table (Table 1) for each infection of interest, with 'X' being the number of cases that would not be captured using any/all three data sources. [Note: 'Yes' means a case was identified in the respective dataset, while 'No' means it was not identified. For example, cell C in the table identifies the number of cases that were identified in encounter records and were also reported but without a positive laboratory test.]

Table 1. Capture-Recapture Methodology Matrix

	Laboratory	Yes	Yes	No	No
	Encounter	Yes	No	Yes	No
Medical Event Reports	Yes	A	B	C	D
	No	E	F	G	X

Case Identification

Case Definition

Case definitions were established using International Classification of Infection, Ninth Edition, Clinical Modification (ICD-9-CM) codes and methods of laboratory isolation described in the Tri-Service Reportable Events Guidelines and Case Definitions (May 2004). Table 2 outlines the criteria for each of the STIs evaluated in this assessment.

Table 2. STI Case Definitions

	Chlamydia	**Gonorrhea**	**Syphilis**
Organism	*Chlamydia trachomatis*	*Neisseria gonorrhoeae*	*Trepenoma pallidum*
ICD-9-CM Codes	099.41	098*	090*, 091*, 095*, 096
Laboratory results	Culture, antigen	Culture, antigen	Culture, antigen, treponemal test

ICD-9-CM codes used to identify STI diagnoses in medical encounters (M2) and medical event reports (DMSS) were defined by the Tri-Service Reportable Events Guidelines and Case Definitions (May 2004). Records were extracted from any encounter where an ICD-9-CM code of interest including all possible extensions was present in any available diagnosis field. A list of all ICD-9-CM codes included in this study and their text definitions are presented in table 3. In the encounter data source, the ICD-9-CM codes are in character format and must be searched by the full term to include only those of interest.

Table 3. ICD-9-CM Codes for Identification of Cases

Chlamydia	Gonorrhea		Syphilis	
099.41	098.0	098.40	090.0	091.61
	098.10	098.41	090.1	091.62
	098.11	098.42	090.2	091.69
	098.12	098.43	090.3	091.7
	098.13	098.49	090.40	091.81
	098.14	098.50	090.41	091.82
	098.15	098.51	090.42	091.89
	098.16	098.52	090.49	091.9
	098.17	098.53	090.5	095.0
	098.19	098.59	090.6	095.1
	098.2	098.6	090.7	095.2
	098.30	098.7	090.9	095.3
	098.31	098.81	091.0	095.4
	098.32	098.82	091.1	095.5
	098.33	098.83	091.2	095.6
	098.34	098.84	091.3	095.7
	098.35	098.85	091.4	095.8
	098.36	098.86	091.50	095.9
	098.37	098.89	091.51	096
	098.39		091.52	

Chemistry and microbiology data were queried for each disease of interest using the search criteria listed in Table 4. The criteria, developed using frequency tables of key fields from chemistry and microbiology data and record review, isolated relevant records and removed any invalid records retrieved. Common invalid records included those indicating quality assurance testing (e.g., lab interoperability processes). All test types available were included for chlamydia and gonorrhea. Valid tests for chlamydia and gonorrhea were identified in chemistry and microbiology laboratory data, and search criteria in Table 4 applied to both data sources. The syphilis search criteria identified three confirmatory syphilis tests: fluorescent treponemal absorption (FTA-ABS) test, *Treponema pallidum* particle agglutination assay (TP-PA), and microhemagglutination assay for *Treponema pallidum* antibodies (MHA-TP). Valid tests for syphilis were identified in chemistry data only. Additional chlamydia and gonorrhea records were removed from analysis due to invalid specimen source and body site collection values (i.e., eye, nail, wound, abscess, etc.).

Table 4. Dataset Search Terms

Type	Chlamydia Terms	Gonorrhea Terms	Syphilis Terms
Inclusion	**Test Name or Test Result field contains one of the following :** CHL TRACH	**Test Name or Test Result field contains one of the following:** GONO GC	**Test Name field contains one of the following:** MHA - TP Fluorescent FTA MHATP TP-PA PARTICLE AGG MIL PUB HEALTH NOTIFICATION **AND Test Ordered or Test Result field contains one of the following:** SYPH PALLID TREPON FTA MHA TP-PA
Exclusion	**Test Name field contains one of the following:** PSITTACI PNEUM IGG IGM AB CHLAMYDO TRACHEA **Test Result field contains one of the following:** CHLAMYDO TRACHEA	**Test Name or Test Result field contains one of the following:** GC/MS GCMS MCG/GCR SEQUENCING X LOGCOP MCG TGC COAG NEG	**Test Name field contains one of the following:** IGM

Data Processing

All three datasets were processed separately, but the methodology was similar. For all datasets, the sponsor Social Security number (SSN) and family member prefix (FMP) variables were formatted as characters to retain leading zeros; records without an SSN were removed. Study data consisted of one CY: 01 January 2006 – 31 December 2006, the last complete year of data at study initiation. Patient category (PatCat) or beneficiary category (BenCat) was used when available (FMP, if not), to include only active duty personnel (PatCat/BenCat = A11, F11, C11, N11, and M11, or FMP = 20). PatCat or BenCat were used when present because active duty service members can be separated from retirees and reservists; FMP values can only separate sponsors from other beneficiaries. PatCat or BenCat were used in the encounter and laboratory databases, while FMP was used for the DMSS database.

Removing Duplicate Records

Duplicate records were removed in the same manner for all datasets. For the purposes of this study, only active duty personnel were included, and therefore the sponsor SSN identified a unique person. Chlamydia and gonorrhea cases were identified using a 30-day gap-in-care rule. Any records that were within 30 days of another record were considered the same case for counting purposes. All syphilis cases were counted once per calendar year due to the inability to determine incident and prevalent cases in the data and the length of infection.

DMSS Records

DMSS data were received in Microsoft Excel format with each infection on its own worksheet. Each infection was imported separately and data were processed as described above. Frequencies were generated to ensure that ICD-9-CM codes included were infection-specific. DMSS records were processed into SAS.

Encounter Records

For the purposes of this study, inpatient and outpatient records were considered the same source, and referred to as "encounter data." Records containing the selected ICD-9-CM codes were extracted from M2 and imported into SAS. Formats of variables were standardized to maintain data quality and allow for subsequent stacking of records. To format and stack the inpatient and outpatient datasets, validated code was used to insert decimals in the ICD-9-CM code field, change variable names to match, and then stack the respective tables.

Records were then searched for ICD-9-CM codes by infection and separated into infection specific datasets. (Note: It was possible for a single encounter to include a diagnosis for more than one infection.)

Laboratory Data

Infection-specific methods were developed to identify cases in chemistry and microbiology laboratory databases. Syphilis testing was identified only in the chemistry database and the case definition differed from chlamydia and gonorrhea, resulting in different analysis methods.

Chemistry

Chemistry records for each of the three diseases of interest were reviewed and classified using information found in key fields. Chlamydia and gonorrhea records were classified as positive or not positive. Syphilis records were classified as suspect or not suspect.

Chlamydia/Gonorrhea:
If non-STI results were present for the records extracted (ex: urine sample for ketones and glucose tests), the record was removed from analysis.

The Test Result field was queried for a positive or negative result and classified accordingly. All test results indicating more information was available in the Result Notes field (e.g., "Comment," "See Note") were reviewed but not included in case classification as coding these records would yield very few additional cases and involve extensive coding. Records whose results status was still unable to be determined were excluded from further analyses. Chlamydia and gonorrhea tests are often completed on the same specimen. Records with test names indicating a test for both infections with non-specific results were not included as a positive case for either disease. Final datasets were created of positive cases for each infection.

Syphilis:
The Test Result field was queried for a positive or negative result and classified accordingly. Records with a non-numeric result (i.e., "Positive," "Negative") were classified using basic text search terms. Records that indicated a reactive, positive, or borderline result were classified as suspect. Records with test results in the form of a titer (e.g., 1:160) were classified as suspect if the test result was greater than 1:80. Records that indicated more information was available in the result notes field were reviewed. If the result notes field indicated a positive confirmatory syphilis test then the test was classified as suspect. All remaining unclassified records were excluded from analyses. A final dataset containing all suspect cases was created.

Microbiology

Microbiology records retrieved for chlamydia and gonorrhea were reviewed and classified as positive or not positive using information found in key fields.

Records were then sorted by SSN, accession number, message date, and set ID (similar to a line number) to observe trends in the data. For both infections, the records were organized such that the first record was the test name and the subsequent record stated the specimen name, if isolated. To include only those records where the organism was isolated, all records with set ID equal to 1 were removed. Subsequently, positive records were identified using search terms for organism names in the test name field. For remaining records, the test result field was queried and classified using basic search terms. All remaining unclassified records were considered not positive. A final dataset containing all positive cases was created.

Matching

Before the databases could be matched, it was necessary to create identifying variables for each source so that the origin of the record could be determined after matching. Variable names and formats were also changed to allow for consistency of matching.

Data were stacked to create one master dataset per infection. Records were then collapsed based on previously defined unique incident case definitions (chlamydia and gonorrhea = 30 days, syphilis = one year). Collapsed records were marked to show in which dataset(s) each case was identified. Frequencies of records in each possible combination of records source(s) were produced.

Results

Cases

The majority (69.9%) of cases were identified in the HL7 formatted laboratory records. There were fewer cases identified in the encounter and DMSS databases with 7.8% and 22.3%, respectively. Of all cases identified in any database, 81.4% were chlamydia (n=45,155), 18.2% were gonorrhea (n=10,080), and 0.4% were syphilis infections (n=220) (Table 5). Note: The cases in each database were not mutually exclusive of each other during this stage of data analysis.

Table 5 illustrates the identified unique incident cases by infection and data source, while Table 6 illustrates the number of identified unique individuals by infection and data source.

Table 5. Number of Unique Cases Identified in each Database by Infection for DOD Active Duty Personnel in CY2006

Infection	Laboratory	Encounters	Medical Event Reports	Total
Chlamydia	32,912	1,908	10,337	45,155
Gonorrhea	5,839	2,301	1,950	10,080
Syphilis	10	127	84	220
Total	**38,761**	**4,336**	**12,371**	**55,455**

Table 6. Number of DOD Active Duty Service Members by Infection in CY2006

Infection	Laboratory	Encounters	Medical Event Reports	Total
Chlamydia	29,840	1,850	10,006	41,695
Gonorrhea	5,474	2,144	1,895	9,506
Syphilis	10	127	84	220
Total	**35,324**	**4,121**	**11,985**	**51,421**

Laboratory Data

There were 32,912 unique cases of chlamydia identified in the HL7 formatted laboratory records for 29,840 active duty DOD service members during CY2006. There were 5,839 unique cases of

gonorrhea identified in the HL7 formatted laboratory records for 5,474 members. There were 10 unique cases of syphilis identified in the HL7 laboratory records for 10 members (Table 5, Table 6). Nine percent (n=3,072) of chlamydia infections were re-infections of the same person; this happened less frequently (n=365, 6.3%) with gonorrhea infections.

Encounter Records

There were 1,908 unique cases of chlamydia identified in the encounter records for 1,850 active duty DOD service members. There were 2,301 unique cases of gonorrhea identified in the encounter records for 2,144 members. There were 127 unique cases of syphilis identified in the encounter records for 127 members (Table 5, Table 6). Re-infection was less frequent in the encounter records. Gonorrhea patients were twice as likely to have re-infection diagnosed as chlamydia patients. There were 58 re-infections of chlamydia (3.0%) and 157 re-infections of gonorrhea (6.8%).

DMSS Records

There were 10,337 unique cases of chlamydia identified in the reported medical event records for 10,006 active duty DOD service members. There were 1,950 unique cases of gonorrhea identified in the reported medical event records for 1,895 members. There were 84 unique cases of syphilis identified in the reported medical event records for 84 members (Table 5, Table 6). Chlamydia had the highest percentage of re-infection, but was much lower and more similar to the gonorrhea re-infections than in the other databases. There were 331 re-infections of chlamydia (3.2%) and 55 re-infections of gonorrhea (2.8%).

Demographic Comparisons

Additional variables present in the databases used for case identification allowed for a comparison of service and gender among data sources for each infection. The extract of DMSS data received did not include gender specific information, though it is available in the master dataset. Infections were considered separately due to the differences in at risk populations and clinical manifestations. Cases may have been identified in more than one dataset and therefore demographic statistics were not mutually exclusive.

Chlamydia

The distribution of service among active duty service members with at least one case of chlamydia in CY2006 was not consistent across the data sources. In the laboratory dataset, the majority of cases were Navy/Marine Corps service members (45.3%) followed closely by the Army (36.5%). In the encounter dataset, the Army and Navy/Marine Corps service members represented approximately the same proportion of cases (42.6% and 41.8%, respectively). The majority of cases reported in DMSS were for Army service members (61.2%), while the Navy/Marine Corps only accounted for 14.1% of cases. The proportions of Air Force service members were consistent among the laboratory and encounter databases (17.6% and 14.5%, respectively), but were higher in the DMSS dataset (24.6%). Coast Guard represented ≤1% in each dataset.

In both the laboratory and encounter databases there were more male than female chlamydia cases. However, the gender proportion was much closer in laboratory than encounter records (Table 7).

Table 7. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Chlamydia Infection in CY2006

	Laboratory (HL7)	Encounter (M2)	Medical Event Reports (DMSS)
Service			
Army	10,899 (36.5%)	788 (42.6%)	6,128 (61.2%)
Air Force	5,256 (17.6%)	268 (14.5%)	2,456 (24.6%)
Coast Guard	176 (0.6%)	20 (1.1%)	8 (0.1%)
Navy/Marine Corps	13,509 (45.3%)	774 (41.8%)	1,414 (14.14%)
Gender			
Male	17,269 (57.9%)	1,438 (77.7%)	-
Female	12,570 (42.1%)	412 (27.3%)	-
Total	**29,840**	**1,850**	**10,006**

Gonorrhea

The distribution of service among active duty service members with at least one case of gonorrhea in CY2006 was not consistent across the data sources. In the laboratory dataset, the majority of cases were Navy/Marine Corps service members (43.9%) followed closely by the Army (42.2%). The majority of cases in encounter records and DMSS were for Army service members (62.1% and 70.5%, respectively), while the Navy/Marine Corps only accounted for 24.4% and 15.2% of cases, respectively. The proportion of Air Force service members was consistent among the all three databases, ranging between 12.8% and 15.2%. Coast Guard represented <1% in each dataset.

In both the laboratory and encounter databases, there were more male than female gonorrhea cases (Table 8).

Table 8. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Gonorrhea Infection in CY2006

	Laboratory (HL7)	Encounter (M2)	Medical Event Reports (DMSS)
Service			
Army	2,311 (42.2%)	1,331 (62.1%)	1,335 (70.5%)
Air Force	743 (13.6%)	275 (12.8%)	287 (15.2%)
Coast Guard	19 (0.4%)	15 (0.7%)	2 (0.1%)
Navy/Marine Corps	2,401 (43.9%)	523 (24.4%)	271 (14.3%)
Gender			
Male	4,081 (74.5%)	1,475 (68.8%)	-
Female	1,393 (24.5%)	669 (31.2%)	-
Total	**5,474**	**2,144**	**1,895**

Syphilis

The distribution of service among active duty service members with at least one case of syphilis in CY2006 was not consistent across the data sources. The distribution was similar in the laboratory and DMSS datasets for all services, with the majority being Army, followed by Navy/Marine Corps and Air Force. In the encounter dataset, the proportion of Army and Navy/Marine Corps was approximately equal, each about 40%. Encounter data was the only source that included syphilis cases from Coast Guard service members.

The gender ratio was approximately the same in both the laboratory and encounter data, but there were disproportionally more male than female syphilis cases (Table 9).

Table 9. Service and Gender Characteristics by Data Source of Patients Diagnosed with at Least One Syphilis Infection in CY2006

	Laboratory (HL7)	Encounter (M2)	Medical Event Reports (DMSS)
Service			
Army	6 (60%)	51 (40.2%)	50 (59.5%)
Air Force	1 (10%)	22 (17.3%)	15 (17.9%)
Coast Guard	0	4 (3.2%)	0
Navy/Marine Corps	3 (30%)	50 (39.4%)	19 (22.6%)
Gender			
Male	8 (80%)	111 (87.4%)	-
Female	2 (20%)	16 (12.6%)	-
Total	**10**	**127**	**84**

Matching

The results of matching the three databases are shown by infection in Tables 10, 11, and 12. Similar characteristics were observed among all infections with most records present in only one database. For chlamydia, there were 395 (1.0%) cases identified in all three databases, 4,725 (11.9%) cases in two databases, and 34,597 (87.1%) in only one database. For gonorrhea, there

were 411 (5.0%) cases recorded in all three databases, 1,153 (14.1%) cases in two databases, and 6,603 (80.9%) in only one database. For syphilis, there were no cases recorded in all three databases, 19 (9.4%) cases in two databases, and 183 (90.6%) in only one database.

Table 10. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Chlamydia

Medical Event Reports	Laboratory	Y	Y	N	N
	Encounter	Y	N	Y	N
	Y	395	3,918	335	5,707
	N	472	28,157	733	

Table 11. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Gonorrhea

Medical Event Reports	Laboratory	Y	Y	N	N
	Encounter	Y	N	Y	N
	Y	411	381	444	714
	N	328	4,723	1,166	

Table 12. Distribution of Matching Records for DOD Active Duty Personnel in CY2006 Infected with Syphilis

Medical Event Reports	Laboratory	Y	Y	N	N
	Encounter	Y	N	Y	N
	Y	0	0	19	65
	N	0	10	108	

Capture-Recapture

The capture-recapture method was applied to the case distribution for syphilis. There were 202 syphilis cases identified (Table 12), but no cases were identified in all three data sources. Also, although ten positive laboratory cases were found, none were reported to DMSS. Of all cases identified, 32.2% were captured only through the DMSS reporting system with no MHS laboratory confirmation or encounter within the calendar year. The capture-recapture method was applied using the Normal distribution assumption, as opposed to the Poisson distribution, because otherwise the zero values where no overlap between databases occurred was considered as null in the generalized model. Under the Normal distribution assumption, the method considers the population to be relatively stable with potential for the infection to be evenly distributed, as well as considering the zero values as actual zero case capture as opposed to null. The model of best fit estimated that approximately 154 cases (95% CI: 147.6, 160.4) may be missing when using the three data sources, for a total estimated case burden of 356. The hierarchical model of best fit consisted of two significant interaction terms, the DMSS to laboratory data interaction (medical event reports*laboratory) and encounter to laboratory data interaction (encounter*laboratory). Therefore, if only the electronic data sources were used as many as 43.3% (N=154/356) of cases many not have been considered for case burden.

Discussion

A major concern with military surveillance models is the inability to capture cases from operational forces and non-MHS facilities. Encounter and laboratory records from these locations are not currently accessible, as they do not utilize the same electronic medical database as fixed MTFs. Operational forces may include, but are not limited to, ships, forward deployed units, and battalion aid stations. All cases can be reported to DMSS or service-specific reporting systems directly, however, under reporting continues to be a problem for many infections and STIs in particular. All DMSS records were considered for this study because the provided records did not distinguish between operational forces and non-operational forces. Through consultation with the Commander Naval Air Forces and other subject matter experts, mechanisms for record keeping aboard ships and at other locations were examined for possible data collection methods. Many operational forces commands do not use uniform or sometimes even electronic record-keeping processes. Some processes are force specific; for example, the Navy fleet forces use the Ship Administration Management System (SAMS) for encounter recording and infection reporting. In the DON, aggregate reports are collected on a regular basis for carriers that detail the number of STIs but do not include personal identifiers. Cases identified in non-fixed medical facilities were not able to be included.

The Centers for Disease Control and Prevention's (CDC) Syphilis Elimination Effort (SEE) has a goal of syphilis eradication in the US, and finding all cases in the DOD would be essential to supporting this initiative. Therefore, each additional case added from any dataset is important. Laboratory tests with no identified related encounter records may need additional follow up to determine the source of testing and if care was received, potentially at a non-MHS facility. The resulting number of estimated cases that would not have been identified in any database was close to half of the overall estimated infection burden. The estimate may be impacted by the lack of alignment between the data, which could skew the results due to the increased variability that each source contributes. There were several data alignments that resulted in no shared cases. The syphilis records were affected highly by the encounter database, as over half of cases identified were found in that database alone. An additional third of cases were only reported to DMSS, which may be impacted by operational forces.

The volume of cases of chlamydia and gonorrhea is much higher than syphilis in the DOD population. Though using all three sources for surveillance may require more resources than single-source surveillance methods, using only one source does not provide a complete picture of the true STI burden. The benefit of continuing use of the passive reporting system is the consistency that may be observed in reporting trends under the assumption that a stable proportion of cases are reported. This project found that a vast majority (81.8%) of chlamydia and gonorrhea cases were identified in the laboratory data, the remainder being found in encounter and/or reporting records. Using encounter or reporting databases alone would have only identified less than a tenth of chlamydia cases and a fourth of gonorrhea cases.

The most significant advantage of using laboratory records for case identification is that laboratory results are often definitive. While this may not be the case with all infections, it is true for STIs. Positive laboratory testing continues to be the gold standard for surveillance

systems and case finding projects. The sensitivity and specificity of these laboratory tests are high. ICD-9-CM codes can be based on many different aspects of a clinical encounter and are influenced by factors including provider impressions, symptom presentation, prioritization of symptoms, misdiagnosis, and miscoding. These records may or may not be updated after laboratory results are received. Also, laboratory tests detect infection in otherwise asymptomatic patients, which may impact initial diagnoses at the clinical encounter. The use of laboratory records is advantageous for surveillance, because it does not require additional reporting measures from providers/technicians and creates a specific case definition where the organism is detected. Reported cases require additional effort on the part of a provider and have been shown to be consistently under reported. This approach to surveillance can be applied to other heavily under reported infections.

Case identification processes were lengthy due to the non-standardized data entry in the MHS HL7 formatted laboratory records. While ICD-9-CM code diagnoses could be used to extract cases from encounter (M2) and reporting (DMSS) datasets, the chemistry and microbiology datasets required queries of multiple free text fields. Microbiology records were more consistent than chemistry records, but chemistry represented more of the cases identified from the laboratory data. Currently, there are efforts underway by the Laboratory Interoperability team to standardize laboratory orders and results entered into MHS databases. Though these efforts would only affect new records entered, they would reduce complexity of analysis in studies performed on the affected data. If future STI analyses are performed for specific commands, the variability of the text fields may decrease due to a higher consistency in recording practices within an MTF. It is not anticipated that the structure of the HL7 format presented any major limitations in the identification of laboratory confirmed STI cases.

Through consultation with a practicing provider at a Navy MTF-based STI clinic, it was possible to address some of the concerns with the lack of alignment among the data sources. Most providers at STI clinics are not physicians, and only physicians can diagnose a patient with an infection, i.e. assign an ICD-9-CM code for a specific infection. The alternative to this is a V-Code diagnosis for STI Counseling, which is not reportable under the current Tri-Service Reportable Guidelines (May 2004), as it may include counseling with or without a diagnosis or testing. The STI clinic providers can request laboratory testing, provide treatment, and have access to service-specific medical event reporting systems. Using this information, it was found that all STI cases identified through encounter record for this project were also associated with a STI counseling V-code diagnosis within 30 days. The STI Counseling V-Code was also present frequently in the encounter records without an STI diagnosis. While this consultation aided with study result interpretation, additional information from other services and types of facilities would allow for a better understanding of STI recording processes. Understanding clinical practice and how ICD-9-CM codes are utilized in the clinical setting is essential to a surveillance system. It would aid in interpreting the distinction between STIs that are and are not reported, help determine the volume of cases that may be excluded due to non-reportable coding, and help determine how to collect this data for policy implementation/program establishment.

Analysis performed for this project identified several topics for possible future studies. When STI cases were identified in the laboratory data, it was noted that many individuals had multiple positive tests more than 30 days apart, indicating a new or re-infection. This project did not assess whether those individuals had a comparably better or worse matching to related encounter records or if the infections were reported to DMSS one or multiple times. Furthermore, the distribution of genders was not consistent among the STIs and did not reflect the gender distribution in the active duty DOD service members. The value of screening programs for females was not evaluated to determine if differences were due to established policies. Both project ideas were submitted as an FY2009 GEIS proposal and accepted.

Overall, the project goals were met. The project was able to establish case definitions in all data sources, extract cases, and align cases between the sources. The project identified the limitations of the capture-recapture method when a particular data source contributes a substantial proportion of all cases identified. It was determined that the Capture-recapture method can be applied after a thorough review of the case alignment. The method was successfully applied to one out of the three infections of interest, with stable statistical results. Results may be able to provide an improved insight into the known infection burden based on existing data sources and the additional estimated burden not captured elsewhere. Furthermore, the project clearly outlined the impact that some data sources have on case estimation and the potential role of laboratory results in the estimation of chlamydia and gonorrhea burden. The methods applied to this project (including case identification) can be confidently applied against other reportable infections that have historically been heavily under reported.

POINT OF CONTACT

Navy and Marine Corps Public Health Center
EpiData Center Department

Communicable Disease Division
757.953.0700

WWW.NMCPHC.MED.NAVY.MIL/
usn.hampton-roads.navmcpublhlthcenpors.list.nmpch-epi@mail.mil

Acronym List

Acronym	Definition
BenCat	Beneficiary category
CDC	Centers for Disease Control and Prevention
CHCS	Composite Health Care System
CY	Calendar year
DMSS	Defense Medical Surveillance System
DOD	Department of Defense
EDC	EpiData Center
FMP	Family member prefix
FTA-ABS	Fluorescent treponemal absorption (confirmatory syphilis test)
FY	Fiscal year
GEIS	Global Emerging Infections Surveillance and Response System
HL7	Health Level 7 (laboratory data format)
ICD-9-CM	International Classification of Infection, Ninth Edition, Clinical Modification
M2	Military Health System Management Analysis and Reporting Tool
MHA-TP	Microhemagglutination assay for *Treponema pallidum* antibodies
MHS	Military Health System
MTF	Military treatment facility
NDRSi	Naval Disease Reporting System- internet version
NMCPHC	Navy and Marine Corps Public Health Center
PatCat	Patient category
SADR	Standard Ambulatory Data Record
SAMS	Ship Administration Management System
SEE	Syphilis Elimination Effort
SIDR	Standard Inpatient Data Record
SSN	Social Security number
STI	Sexually transmitted infection
TP-PA	*Treponema pallidum* particle agglutination assay